Tail protection for long investors: Convexity at work


May 9, 2016

Abstract

We relate the performance of trend following strategy to the difference between a long-term and a short-term variance. We show that this result is rather general, and holds for various definitions of the trend. We use this result to explain the positive convexity property of CTA performance and show that it is a much stronger effect than initially thought. This result also enable us to highlight interesting connections with Risk Parity portfolio.

Finally, we propose a new portfolio of options that gives us a pure exposure to the variance of the underlying, shedding some light on the link between trend and volatility, and also helping us understanding the exact role of hedging.
# Contents

1. **Introduction**  
   - 1.1 The trend on a single asset  
2.  
   - 2.1 A toy model for the trend  
   - 2.2 Trend following using an exponential moving average  
   - 2.3 Changing the shape of the signal  
   - 2.4 A word on skewness  
   - 2.5 Discussion  
3. **Convexity at work on real data**  
   - 3.1 Understanding the SG CTA Index  
   - 3.1.1 The pool  
   - 3.1.2 Comparing the performances with the SG CTA Index  
   - 3.1.3 The convexity of the SG CTA Index with the stock market  
   - 3.2 Convexity and diversification  
   - 3.2.1 Estimating the convexity of a diversified trend with respect to the stock market  
   - 3.2.2 A global portfolio  
   - 3.2.3 Reproducing a Risk-Parity index  
   - 3.2.4 Link between Trend and Risk Parity  
4. **The link with options**  
   - 4.1 A collection of strangles  
   - 4.2 Trend and re-hedging  
   - 4.3 Conclusions  
5. **Summary and perspectives**  
6. **Appendix**  
   - 6.1 Proof of trend formula  
   - 6.1.1 Theorem  
   - 6.1.2 Discrete equation for EMA filter  
   - 6.2 Generalized trend formula for trend-following  
   - 6.2.1 Trend formula in continuous-time framework  
   - 6.2.2 Linear trend estimation  
   - 6.2.3 Non-linear trend estimation  
   - 6.3 Effect of diversification  
   - 6.3.1 Convexity versus index  
   - 6.3.2 Convexity versus risk parity  
   - 6.4 Strangle portfolio: Mark-to-the-market and Greeks
1 Introduction

A key concept in finance is the idea of diversification: if you can find decorrelated alphas, your total performance will be smoother, and your Sharpe ratio higher (see [2] for the original paper). Even better is the situation when one strategy registers large gains on average when others take a hit (negative correlation). In particular, since a lot of investors have a long stock exposure, a strategy that performs well (or at least does not turn south) when the market goes down sounds like a very useful diversification, and should be actively sought by long investors.

A typical example of such a strategy is a long-option portfolio. Specific option portfolios, called variance swaps, are designed to give one an exposure to the realized variance of an asset, therefore providing direct protection when the volatility rises (see for example [20, 21] for detailed reviews on the topic). Unfortunately, these portfolio typically exhibit strong negative drifts, which makes them a costly protection indeed.

Hedge funds are yet another alternative. The hedge fund industry has always claimed that its performance is not correlated to the market, and therefore hedge funds should be viewed as a source of diversification, as well as providing pure alpha. However, if we plot the monthly performance of the HFRIx index, a global Hedge Fund index, as a function of market return, like we do in fig.1 we see a strong positive correlation between the two variables (i.e. when the market goes down, so do hedge funds returns), and a negative convexity: the performance is worse in periods of high market volatility, so that if we try a quadratic fit, the leading coefficient is negative. In other words, hedge funds seem to have difficulty fulfilling their promises in terms of diversification (see [2] for a detailed analysis of hedge fund performance).

One interesting exception to this is the case of CTAs: in 2008 in particular, their performance during the Lehman crisis was very strong, triggering a subsequent massive rise in assets (current estimates are in the range of USD 300 bns, see [12]). This has been confirmed by various studies who have looked at CTA performance during or immediately after market crashes, and found above average performances [2, 10] (followed by below-par results, see [11] and references therein). We reproduce the basis of these conclusions in fig.2 and 3 where we plot monthly returns of the two main CTA indices as a function of market performance. As we can see, these strategies have indeed performed on average better when the market volatility has been high.

Unfortunately, this convexity is based on flimsy statistical evidence (see [4] though), as quantified by the very low value of the $R^2$ in both fits (around 0.02). Therefore, it seems crucial to make sure that we are not looking at a statistical artifact. We will try in this paper to understand the mechanism at work behind this convex behaviour, and to find a better way to quantify it, so that we can be assured this is not a statistical fluke (which we confirm in fig.2). To do this, we will need to understand the trading patterns by CTAs, and the strategies they use.

As already noticed [2] (and we will demonstrate it again in this paper), CTA performance can actually be explained by one very simple strategy: trend following [13]. This strategy's robustness and stability to a wide range of parameters, as well as the fact that it works across a large set of asset classes, make the out-of-sample performance even more impressive, while the simplicity of the idea makes it very suitable to the algorithmic trading style favoured by many CTA funds [14]. In a previous paper, we have already studied the significance of this strategy, and found that it is persistent over 2 centuries [16], making it one of the most significant market anomalies ever documented (barring some HF effect). The problem we therefore want to tackle now is to understand, and quantify, the convexity inherent to any trend following strategy.

We start by deriving some equations relating the performance of a single-asset trending strategy to the difference between a long-term and a short-term variance. We show that this result is rather general, and holds for various definitions of the trend. We can then prove that this single-asset trend shows the expected convexity properties. To understand the CTA performance, we replicate the SG CTA Index, using surprisingly few parameters. We can then understand how to measure the convexity, and we will see a much stronger effect than naively plotted in fig.2. We will then investigate in detail the impact of diversification on this convexity, and find interesting connections with Risk Parity strategies. Finally, we look at the most simple convex strategy: buying options. We consider a new portfolio of options that gives us a pure exposure to the variance of the underlying. We find that the pay-off of these simplified variance swaps appears to be strikingly similar to that of a simple trending strategy. This shed some light on the link between trend and volatility, somewhat justifying the "long vol" attribute of trend strategies, but also helps us understand the differences between long-option portfolios and trending, and the exact role of hedging.
Figure 1: Monthly returns of HFRI index vs. monthly returns of S&P. We see a concave relationship: hedge funds do worse when volatility is high.

Figure 2: Monthly returns of BTOP 50 vs. monthly returns of S&P. We fit this with the ansatz $y = ax^2 + bx + c$. $R^2 = 0.02$. The convexity is visible, but the $R^2$ is low.

Figure 3: Monthly returns of SG CTA Index vs. monthly returns of S&P. We fit this with the ansatz $y = ax^2 + bx + c$ and get a convex function ($R^2 = 0.02$).
2 The trend on a single asset

2.1 A toy model for the trend

We begin this study with a simplified model, which will make all the derivations very straightforward, while keeping the main features of the more sophisticated models. We consider a predictor on day $t$ whose value is simply the price difference between $t$ and $0$:

$$P_t = S_t - S_0 = \sum_{i=0}^{t} \Delta_i$$

where $S_t$ is the price of our asset at day $t$, and $\Delta_t = S_t - S_{t-1}$ is the price difference (or return) between days $t-1$ and $t$.

The daily P&L from $t-1$ to $t$ will be:

$$G_t = \Delta_t \Delta_{t-1} \sum_{i=1}^{t} \Delta_i$$

Now we can aggregate the performance of our trending predictor from day 0 to a given day $\tau$, and rearrange the sums:

$$\sum_{t=2}^{\tau} G_t = \sum_{t=2}^{\tau} \sum_{i=1}^{t-1} \Delta_i \Delta_i = \sum_{2 \leq t, i < t} \Delta_t \Delta_i$$

$$2 \sum_{t=1}^{\tau} G_t = \left( \sum_{i=1}^{\tau} \Delta_i \right)^2 - \sum_{i=1}^{\tau} \Delta_i^2 = (S_\tau - S_0)^2 - \sum_{i=1}^{\tau} \Delta_i^2$$

This simple formula is at the core of our understanding of the performance of a trending system. In a nutshell, it says that the performance, aggregated over $\tau$ days, is proportional to the difference between the realized variance computed over $\tau$-days and the variance computed using 1-day returns. It makes perfect sense, since the difference between these 2 terms is the sum of cross-terms $\Delta_i \Delta_{i'}$, which are related to the auto-correlation of the return. So the trend performance is positive if returns are auto-correlated, which is very intuitive.

This calls for two more comments. First, we considered our time unit to be a day, but really, it is the time over which we change our predictor, and rebalance our portfolio: this formula applies to a fast, intraday trend as well as a monthly trend. We simply need to compute the short-term variance over 1 time step, whatever that time step is.

Secondly, we have considered here $\Delta_t$ to be price differences. Following simple optimization procedures to manage the heteroskedasticity of financial time series, it is quite customary to normalize these returns by an estimate of the volatility, to have objects with unit variance (or as close to that as possible). A typical example is some past exponential moving average of realized volatility over time-scale $\tau$:

$$\sigma_t = a \sqrt{\mathcal{L}_\tau[\Delta_i^2]}$$

with $\mathcal{L}_\tau[X_i] = (1 - \lambda) \sum_{i \leq \tau} \lambda^{i-1} X_i$, and $\lambda = \frac{2}{\tau+1}$. $a$ is an empirical factor that takes into account the fat tails of the return distribution, as well as the correlation of volatility. In practice, we considered $a = 1.05$ and $\tau = 10$ in the rest of this paper.

We can reproduce all the computations above with normalized returns $R_t = \frac{\Delta_t}{\tau_{t-1}}$, and find that the performance of this new, risk-managed trend aggregated over $\tau$ days is:

$$\sum_{t=1}^{\tau} G_t = \frac{1}{2} \left( \left( \sum_{t=0}^{\tau} R_t \right)^2 - \sum_{t=0}^{\tau} R_t^2 \right)$$

We notice that, by definition, $< R_t^2 > = 1$ if our volatility estimator is unbiased so the last term of the equation is known on average. This means that the performance of this simple trend is simply governed by the long-term variance of the underlying.
Figure 4: $M_{\tau}[G]$ as a function of $P$, the past trend, for a trending system on the S&P with $\tau = 180$.

If we consider a long history of this strategy, and split it in sub-samples of size $\tau$, we can plot the performance aggregated over this time scale as a function of the (normalized) return of the underlying over the same time scale. This return is precisely $X = \frac{1}{\tau} \sum_{i=1}^{\tau} R_t$. We have already pointed out that the second term in the formula above averages to 1, so we get the conditional expectation of the pnl:

$$\langle G \rangle_X = \frac{\tau}{2} (X^2 - 1)$$

This formula is very simple, and clearly shows that we should expect convexity from such a simple strategy, but only over a certain time-scale: it should work well when the underlying asset has experienced large moves over the trending time horizon. On the other hand, it does clearly not provide a protection for a massive movement that happens over 1 time step; any price "jump" is not hedged by a trending strategy, and its average performance should be null.

2.2 Trend following using an exponential moving average

In reality, most funds use moving averages to compute their trending signals [18]. Therefore, it is important to generalize and extend our formula above to these types of filters. In this subsection, we consider a very standard trend-following strategy, which uses the exponential moving average (EMA) filter to define its prediction $P_t = \sqrt{\tau} L_{\tau}[R_t]$ (the extra factor $\sqrt{\tau}$ simply ensures that our predictor has unit variance).

Defining $\tau' = \frac{1}{2} + \frac{1}{2\tau}$, and $G_t = P_{t-1} R_t$, we can write (see Appendix 6.1.2 for the proof):

$$L_{\tau'}[G_t] = \frac{\sqrt{\tau}}{\tau - 1} (P_t^2 - L_{\tau'}[R_t^2]) \quad (1)$$

As we can see, here as well, the performance, once averaged over a suitable period of time, can be rewritten as a difference between 2 variance-type expressions: the first based on the aggregated past return (i.e. the position of our trend following system), and the second based on the 1-day returns properly re-normalized. Rather than the average performance, we may want to consider $M_{\tau'}[G_t] = \tau' L_{\tau'}[G_t]$, which is closer to the aggregated performance over time $\tau'$. If our volatility estimate is un-biased and ensures that the second variance stays close to 1, then we have

$$\langle M_{\tau'}[G] \rangle_P = \frac{\tau' \sqrt{\tau}}{\tau - 1} (P^2 - 1) \quad (2)$$

As an illustration of the validity of this formula, we have plotted in Fig.4 the cumulated performance over $\tau'$ days $M_{\tau'}[G_t] = \tau' L_{\tau'}[G_t]$ for $\tau = 180$ of a trend strategy with target volatility around 15%. As we can see, we get a very good agreement between the theoretical line and the realized performance. In particular, the convexity is quite visible in this set-up, while it would have been utterly blurred had we looked at 1-day, or even 1-month, returns.

This sensitivity to the averaging time means that we need to estimate carefully the time-scale CTAs use to define their trend if we want to see any clear sign of convexity in the CTA performances.

2.3 Changing the shape of the signal

In practice, people use a wide range of filters to capture the trend anomaly: square average instead of exponential, Vertical-Horizontal Filters, Strength Indices, crossing of 2 price averages... We want
to present here a case of particular interest: what happens if we simply take the sign of our signal, rather than scaling our positions with the strength of it? This is quite common in practice, since it prevents us from taking extreme positions on one single asset, and loose the benefits of diversification.

We can show that, in this case, we can rewrite the performance in the following form (see 6.2.3):

\[
\langle M_\tau[G_t] \rangle_P = \sqrt{\tau} \left( |P| - \sqrt{\frac{2}{\pi}} \right). \tag{3}
\]

We have plotted in fig. the typical shape of the cumulated performance over \( \tau \) days \( M_\tau[G] \) as a function of \( P \), and here as well, we get the best performance when the market experiences large volatility over the time scale used to compute the trend. Instead of a parabolic fit, however, we see a piece-wise linear profile.

Since even in this extreme regime, we get a convex behaviour, we believe that intermediate cases where our positions should interpolate between the parabola of eq.1 and the V-shaped curve we obtained here, while keeping the convex feature of the trend.

### 2.4 A word on skewness

Skewness has been proposed as a measure of risk when determining what is a risk premium (see [17] for a thorough discussion on the subject). It is usually assumed that the skewness of trend following is positive (see [14], and [2] for a general framework). Here, we can see that this follows directly from eq.1 on long time-scales at least. Indeed, if we assume that the returns follow a gaussian distribution, then so does \( P_t \), which is a linear combination of gaussian returns. Finally, the aggregated P&L varies as the square of \( P_t \), which means that it follows a \( \chi^2 \) distribution, which has a known positive skewness. So we see that the aggregated P&L of a trend does structurally show positive skewness.

On a daily time scale, this is more subtle. If the returns are well-behaved, and exhibit some level of auto-correlation (i.e. if the P&L of the trend is positive on average), then it should exhibit positive skewness. This explains the measured skewness on daily time-scales, since trend following has been a winning strategy over the last decades. In the absence of auto-correlation, however, there is no reason to measure a positive skewness, as was already pointed out in [1].

### 2.5 Discussion

We have seen that the performance of a trending strategy, once aggregated over a suitable time horizon, can be rewritten as a difference between a long-term and a short-term variance. If we properly risk-manage our portfolio, the short-term variance is on average constant, and therefore the P&L is directly linked to the square of the past long-term return. In other words, there is a strong convexity on this single-asset strategy.

It is striking that this conclusion does not depend on the overall performance of the trend itself; we may assume returns to be perfectly decorrelated, and hence the trend to have vanishing expected return. Nonetheless, there would be convexity if we aggregate returns on the proper time scale.
Figure 6: Liquid futures per sector. This selection is stable across time, and does not vary for decades.

We have also established that this result does not depend on the exact shape of the filter: we go from a parabola to a V-shaped function if we take the sign of the average return, but the idea remains the same.

It is also interesting to recall that we are limited by the re-balancing period of the trend: if we take an N-day trend, re-computed every n days, we can expect some protection when the market moves a lot over N days, but there is no such thing if there is a large sudden movement on a period of time smaller than n days; our system is just blind to these oscillations. We need a quicker trend to take these into account.

3 Convexity at work on real data

In this section, we want to understand how our model is related to real data based on CTA performance. In particular, we consider the SG CTA Index\footnote{Data available at the following URL: http://www.barclayhedge.com/research/indices/calyon/} and show that our simple trend allow us to reproduce its performance (as previously shown in \[13\]). We find that the convexity we measure once we get the right time scale, though much larger than what we measure in the introduction (see fig\[3\]), is not as high as what we observed in the previous section. This is of course because of the diversification of contracts over which CTAs operate. Therefore, we will link the convexity of the diversified trend with the performance of a diversified long portfolio, which we find to be very close to that of a Risk Parity index.

3.1 Understanding the SG CTA Index

3.1.1 The pool

We first want to compare the performance of a simple trend following system to that of the most famous CTA index: the SG CTA Index. For simplicity, we considered in our simulations only futures contracts in 5 sectors: stock indices, government bonds, short rates, foreign exchange, and commodities. In each of these sectors, we considered the most liquid contracts. We end up with a selection outlined in table 6.

We believe that this selection is rather un-biased, since liquidity is quite stable over time, and these contracts have been available for at least a few decades. In any case, adding or removing a few contracts should not affect our results.

3.1.2 Comparing the performances with the SG CTA Index

We have already defined a predictor $P_i$ which has unit variance and is based on an EMA of the past returns. We also have seen that we want a position on each asset which scales with the inverse of the volatility so as to achieve a gain in percent $\frac{G_t}{\sigma_t}$ on asset $k$ with constant volatility (to first order). We now want to be more specific: we define $\Pi_k^t = \frac{P_k^t A}{\sigma_k^t}$ to be the position a typical CTA will take to follow predictor $P_k^t$ on asset $k$. $A$ is the assets under management of the fund, and $L$ is its overall leverage ratio, used to achieve the relevant target volatility. These 2 quantities are constant across time and assets. We also call $w_k$ the weight allocated to the strategy on asset $k$, so that we can rewrite the total gain in percent as:

$$G_t = \sum_k w_k G_t^k = \sum_k w_k \frac{\Pi_k^{t-1} \Delta_k^t}{A} = L \sum_k w_k P_k^{t-1} \frac{\Delta_k^t}{\sigma_k^{t-1}},$$

which is what we were looking for. We see that the total gain is the sum over all assets of terms with unit volatility, multiplied by the weight $w_k$. 

$$\frac{\Delta_k^t}{\sigma_k^{t-1}}$$
Figure 7: correlation between $\tilde{G}_t$ and the SG CTA Index as a function of the time-scale of the trend we used $\tau$. As we can see, the maximum is around $\tau = 180$ days.

Figure 8: Cumulated returns of $\tilde{G}_t$ and the SG CTA Index. We seem to capture all the alpha contained in the SG CTA Index with our simple replicator.

To compare our simulated performance $G_t$ with the SG CTA Index, we need to determine the value of $L$, and determine the weights $w_k$. Once again for the sake of simplicity, we choose to allocate the same amount of risk on each asset: $w_k = 1/N$. We believe we could probably achieve a higher correlation with a more sophisticated risk allocation (like in [15] for example), but we want to keep this procedure as simple as possible, and, as we will see, this should be enough to allow us to determine the time-scale of the trend used by CTAs. We then determine the value of $L$ by requiring that our simulated P&L should have the same volatility than the SG CTA Index. We find that we need $L = 0.01$ no matter the time-scale of our trend, as long as $P^A_t$ is properly normalized.

To be really accurate, and compare this strategy with traded funds, we need to make some assumptions on the fee structure. We assume a 1% flat fee for transaction costs $c_t$, and a typical 2% management fee-20% incentive fee for the fund $f_t$. We also took the 3-6 Month Treasury Bill index as the risk-free rate $r_t$, and define $\tilde{G}_t = G_t - c_t - f_t + r_t$ as the total virtual performance of a typical fund net of all fees and including capital gains.

To determine the relevant time scale, we can now compute $\tilde{G}_t$ and correlate this with the SG CTA Index returns: $G_t$. The result of this procedure is plotted in fig.8. As we can see, there is a maximum in the correlation around $\tau = 180$. This value is the typical time scale of the CTA trend.

What is actually striking is how high the correlation is (above 80%), while all we did was to perform a basic, un-sophisticated trend on the most liquid assets in the planet. Fig.8 shows that we actually capture most of the alpha contained in this index, since the Sharpe ratios of the two strategies are very close. This shows that some of the alpha captured by CTAs can be understood in terms of a very simple strategy.

3.1.3 The convexity of the SG CTA Index with the stock market

We have seen that we can reproduce quite accurately the performance of CTAs as measured by the SG CTA Index, using a simple trend with very few parameters. We only need one global multiplier
Figure 9: Cumulated performance over \( \tau' \) days of the SG CTA Index as a function of \( P_t \) on the S&P future, with \( \tau = 180 \). \( R^2 = 0.18 \), which is quite better than what we have if we use monthly data (see fig.2).

\( L \) to ensure that our simulation has the right volatility, and we fit the time scale of the trend to maximize the correlation. We find the optimal value \( \tau = 180 \), and achieve a very good level of correlation for such a crude simulation, above 80%. No doubt that we could improve that if we allow for a more sophisticated portfolio allocation.

This is beyond the scope of this paper, however: we were merely looking for the best value for \( \tau \), which we find to be \( \tau = 180 \), and \( \tau' = 90 \). We can now see if the convexity is in any way more apparent if we average like we propose in Eq.2. We therefore compute \( M_{\tau'}[G_t] = \tau' \mathcal{L}_{\tau'}[G_t] \), the cumulated SG CTA Index performance over \( \tau' \) days, as a function of the S&P return over the trending time scale \( \tau' P_t = \sqrt{\tau} \mathcal{L}_{\tau}[R_t] \). The result is plotted in fig.9 and as we can see, we do get a much more convincing \( R^2 \) (0.2 instead of 0.02).

So it does seem there is some convexity in the SG CTA Index, if we average our variables over the right time scales. The still relatively low value of the \( R^2 \) can intuitively be understood: it comes from the fact that we need more than one asset to reproduce the index, and therefore, we only have a noisy version of the single-asset convexity we studied in the previous section. We therefore need to understand how this diversification will affect the convexity of a trending strategy.

### 3.2 Convexity and diversification

We have seen in the previous section that the convexity of the SG CTA Index with respect to one single asset, though significant, is far from perfect because of its diversification. Therefore, we should estimate how this diversification affects the relationship outlined in eq.2. To further understand what portfolio a diversified trend is actually hedging, we will introduce a long-only diversified portfolio, and show that a diversified trending strategy is convex with respect to this product. We finally link this long-only portfolio to Risk Parity indices.

#### 3.2.1 Estimating the convexity of a diversified trend with respect to the stock market

As explained above, convexity in the industry is often understood in terms of stock market returns: what does a strategy do if the stock market moves a lot, either up or down? That is perfectly understandable since many investors are exposed to the stock market, and are therefore eager to have a hedge in case of a market crash. We have seen before that a trending strategy on the index itself provides a strong hedge, but this leads to rather poor performances. CTAs usually propose a diversified strategy, which therefore dilutes the convexity.

We now want to quantify this. In Appendix 2, we show that, in a certain limit (negligible drifts in all assets, correct normalization of returns...), we can rewrite the relationship between the cumulated return over \( \tau' \) of a diversified trend strategy and the past trend on the stock index as:

\[
M_{\tau'}[G_t] \propto \bar{\beta}^2 (P_t^2 - 1)
\]

with \( P_t \) the trend on the index, and \( \bar{\beta}^2 = \sum_k \beta_k^2 \) the average of the squared \( \beta \) between each asset and the stock market.
This result is rather intuitive: if the market moves a lot over \( \tau' \), showing a high return autocorrelation, then an asset with either a very positive or a very negative \( \beta \) will show a similar amount of autocorrelation (i.e. strong performance for a trending strategy), so it makes sense to consider \( \beta_k^2 \). Alternatively, when \( \beta_k \) is close to 0, there is no reason to expect any hedge from this asset in case of market turmoil.

Eq.5 explains why our \( R^2 \) is not higher in Fig.2: all assets are obviously not 100% correlated with the S&P! Incidentally, the correlation we get between \( \mathcal{M}_t'[G_t] \), the aggregated performance, and \( P_t^2 \) on the S&P is a measure of the average \( \beta^2 \) of our portfolio with the S&P. With the portfolio outlined above, and the value of our \( R^2 \), we get \( \beta^2 = 0.23 \), which seems consistent with the pool of assets we consider.

### 3.2.2 A global portfolio

We now want to introduce a portfolio in which we have a long position on every asset that we use in our SG CTA Index replicator. Namely, we consider positions \( A \frac{\Delta_k}{\sigma_{k-1}} \), where \( A \) represents the assets under management, and \( L \) the leverage we use, so that the percentage returns of this portfolio can be written as:

\[
G_t^{LP} = L \sum_k w_k \frac{\Delta_k}{\sigma_{k-1}}
\]

This portfolio is some sort of a global portfolio, being exposed to everything we have in our universe: indices, bonds, commodities, and a basket of currencies. We also note that the risk taken on every asset is the same.

Now, to hedge this portfolio, we could use a trend following strategy on its own performance \( G_t^{LP} \): we go long when this portfolio has made money over the last \( \tau \) days, and short otherwise. That would lead to results very similar to the ones shown in Fig.1, namely, we would find a very strong convexity.

We now want to show that the diversified trend CTAs do also provide a hedge for this portfolio, that is actually better than just a trend on the global portfolio. To do so, and remembering that \( G_t \) is the P&L of the SG CTA Index replicator, we notice that:

\[
\langle \mathcal{M}_t'[G_t] \rangle_{\tilde{P}} = \sum_k w_k \langle \mathcal{M}_t'[G_t^k] \rangle_{\tilde{P}} = \sum_k w_k L \frac{\tau' \sqrt{\tau}}{\tau - 1} ((P_t^k)^2 - 1) \geq L \frac{\tau' \sqrt{\tau}}{\tau - 1} (\tilde{P}^2 - 1)
\]

where \( \tilde{P} = \sum_k w_k P_t^k \), and we used the convexity of the parabola in the final step of the derivation.

Using Eq.6, we can rewrite this in the form:

\[
\mathcal{M}_t'[G_t] \geq \frac{\tau' \sqrt{\tau}}{\tau - 1} \left( (\mathcal{M}_t'[G_t^{LP}])^2 - 1 \right)
\]

Eq.7 shows how good the diversified hedge performs: while a normal trend on the portfolio ensures that the performance stays on average on the parabola, the diversification means that we are on average above this parabola. This means that CTAs do provide a very significant protection to large moves in this global portfolio we have built. We now have to find out how we can interpret this long-only portfolio.

### 3.2.3 Reproducing a Risk-Parity index

Risk-Parity is a class of investment that, like CTAs, takes positions in a wide range of securities. The main feature of this strategy is, as emphasized by its name, the fact that each asset class provides the same contribution to the global risk (Equal Risk Contribution, see [19] for a excellent introduction on the subject). To first order, assuming the asset cross-correlations all take the same value, it means that the positions on a given asset scales inversely with its volatility, just like our positions when we reproduce the SG CTA Index. In Risk-Parity, however, we only take long positions, which makes it a good candidate to understand our global portfolio.

Based on these similarities, we tried to compare our portfolio with J.P. Morgan Risk-Parity index using the same idea than for the SG CTA Index. Here as well, we fit \( L \) in Eq.6 so that our simulation has the same volatility than the real index, (we find \( L = 0.0114 \)), and we use equal weighting amongst our pool. We also made assumptions for the costs: no execution costs (always long), and a 1% flat management fee. The risk-free rate is the same, so that we can also define \( G_t^{RP} = G_t^{LP} - f_t^{RP} + r_t \) as the performance of our global portfolio, net of all fees.

Based on these assumptions, we are able to reproduce the J.P. Morgan Risk Parity index with a very good precision (see fig.10): the correlation is as high as 89%! No doubt that here as well,
we could increase this level by choosing a more appropriate set of product weights through a more sophisticated portfolio construction, but we feel this is good enough for our purposes, and it would not affect our results anyway.

3.2.4 Link between Trend and Risk Parity

We can now understand the meaning of our global portfolio, that is actually very close to a Risk Parity investment. Therefore, Eq.7 tells us that a diversified trend provides a very good hedge for a Risk Parity investor.

To confirm this, we have used our proxies for CTA and Risk-Parity performance $G_t$ and $G^{T}_{t}$. In fig.[11] we plot $M_t[G_t]$ as a function of $M_t[G^{T}_{t}]$. As we can see, the inequality is indeed satisfied. We reproduced this experiment, using real data from the official indices (we needed to put back the costs to make the performances comparable, however, so we used the assumptions we made in the previous section). We show the result in fig.[12]. In both cases, the inequality is very well satisfied.

What this means is that trend following does naturally provide a hedge to those investors exposed to Risk Parity. Indeed, as eq[7] shows, when Risk Parity experiences a large drawdown, we can expect a good performance from the trend.

4 The link with options

We now turn to another feature of trend following strategies: various papers have mentioned their similarities with long-option portfolios, such as At-The-Money (ATM) straddles (a portfolio composed of a Put and a Call option both at the money). This is a very natural connection, since both strategies have known convex properties. We want to link these results to our formalism. We will propose a new option portfolio that allow us to capture a long-volatility performance in a very
simple, model-free way (that does not require any black-Scholes inspired machinery). The pay-off of this portfolio is actually very close to that of a trend following strategy. We will see how a trending strategy is very naturally interpreted as the hedge of this portfolio, and how this hedge essentially swaps the long-term variance for the short-term one. It also gives us a more detailed understanding of the volatility we are exposed to in each case, and why the price we pay for the protection is different.

4.1 A collection of strangles

We begin our study by considering a simple straddle, composed of an ATM Put and an ATM Call option (similar to what was considered in [23]). It is quite straightforward to verify that the pay-off of this portfolio can be written as:

\[ G_{[0,T]} = |S_T - S_0| - (C_{S_0,T} + P_{S_0,T}) \]

where \( \sigma_{atm} \) is the ATM implied volatility. We notice that we are exposed to the total return between 0 and T. To get an expression which looks closer to eq.1, we need to consider an infinite equi-weighted collection of strangles, centered at the money:

\[
\text{Uniform profile} = 2 \int_0^{S_0} (K - S_T) dK + 2 \int_{S_0}^{\infty} (S_T - K) dK = (S_T - S_0)^2
\]

After a simple integration, we can see that this portfolio has the following pay-off:

\[ P_{\text{Strangles}}^{[0,T]} = \frac{1}{2} (\Delta_{[0,T]}^2 - T \bar{\sigma}^2) \tag{8} \]

where \( \bar{\sigma} \) is defined by:

\[ \bar{\sigma} = \sqrt{\frac{2}{T} \left( \int_0^{S_0} P_{K,T} + \int_{S_0}^{\infty} C_{K,T} \right)} \]

Eq.8 is quite suggestive. It tells us that this portfolio receives the long-term variance \( \Delta_{[0,T]}^2 \), and pays a fixed price at the start of the trade: \( \bar{\sigma}^2 \). Note that this long-term variance is the same we receive when we follow our toy trend following model at the start of this paper. The only difference is the price we pay to receive this variance: realized vs. implied volatility. Since options are notoriously sold at a premium, an option strategy would be more expensive than a trend following one.

In practice, we can only buy a finite set of options, but, just like any trend strategy should interpolate between the linear and the fully capped trend (Eq.1 and Eq.5), we believe that the pay-off of our real portfolio should lie somewhere in between the pay-off of our infinite collection of strangles, and that of a single straddle. The farther we can go, the larger the quadratic region will be.
4.2 Trend and re-hedging

At \( t = 0 \), our option portfolio satisfies \( \Delta = 0 \), since we consider strangles centered at the money. We can try intuitively to understand what is the hedging trade we have to do when the spot price moves from \( S_0 \) to \( S_t \). Figuratively, we have to re-center our portfolio \( P \) around the new price \( S_t \), so we have to exchange Calls for Puts (Puts for Calls) if the price has gone up (down) between \( S_t \) and \( S_0 \). But selling a Call and buying a Put at the same strike is the same as selling the underlying, so the hedging tells us to sell the underlying, by an amount \( S_t - S_0 \) when the price goes up, and to buy it if it goes down (see fig.13 for an illustration of this procedure).

This trade hedge is exactly the opposite of what our toy-model for trend following does, so we know the hedge performance can be written as:

\[
Pnl_{hedge} = -\left(\Delta^2_{[0,T]} - \sum_t \Delta_t^2 \right)
\]

We derive in Appendix 6.4 formulae for all the usual greeks in a more mathematical way, but it is quite neat to see that we can get a feeling for \( \Delta \) that is accurate.

If we add the pay-off of our portfolio with this re-hedge performance, we get the total performance of our hedged portfolio of strangles:

\[
Pnl_{[0,T]}^{Hedge,Str} = \sum_t \Delta_t^2 - T \bar{\sigma}^2
\]  

4.3 Conclusions

As we can see, the hedging strategy exchanges the variance of the total return over the period \([0, T]\) with the variance defined using the re-hedging frequency (see [22] for a detailed account on the role of hedging, and [24] for a derivation of the importance of the hedging frequency in a different set-up). At the end, our portfolio exhibits the exact performance of a variance swap, but the formulae we used are model-independant, we never need Black-Scholes inspired models. This type of construction also helps us understand that it is the hedging strategy that decides which volatility we are exposed to: if we re-hedge every 5 minutes, we get the volatility of 5-minutes returns, while we get daily volatility if we re-hedge at the close of the day. The price we pay for the options \( \bar{\sigma} \), on the other hand, is always the same.

It also helps us clarify the similarities between a trend following system and a long naked straddle portfolio. They are both exposed to the long term variance \( \Delta^2_{[0,T]} \) (i.e. they both provide protection for a long-stock portfolio), but the second one buys that exposure at a fixed price \( \bar{\sigma} \) while the first one pays the short-term variance \( \sum_t \Delta_t^2 \). An interesting consequence is that, if a shock occurs in one single day that dominates the return over the interval \([0, T]\), the naked straddle will make money, since the entry price is fixed, while the trend will be flat on average. In other words, the straddle is a better hedge, and therefore its price \( \bar{\sigma} \) should be higher than the real volatility \( \sum_t \Delta_t^2 \).

The premium we pay on options is however quite high, and long-gamma portfolios have been consistently losing money over the past 2 decades, while trend following has actually posted positive performance. So, even if options provide a better hedge, it still seems that trend following is a much cheaper way to hedge a long-only exposure.
5 Summary and perspectives

In this paper, we have shown that a single-asset trend has a built-in convexity if we aggregate its returns over the right time-scale. This becomes apparent if we rewrite the performance of the trend as a swap between the variance defined over long-term returns (typically the time scale of the trending filter) and the one defined over short-term returns (the rebalancing of our portfolio). This feature appears to hold for various filters and saturation levels.

The importance of these 2 timescales has been underlined, and it is clear that the convexity (and the hedging properties) are only present over long-term time scales (as defined by the trending filter itself): it is wrong to expect a 6-month trending system rebalanced every week to hedge against a market crash that lasted only a few days.

We also turned our attention to CTA indices, and particularly the SG CTA Index. We have proposed a simple replication index, using a very natural un-saturated trend on a pool of very liquid assets. Assuming realistic fees, and fitting only the time-scale of the filter, we get a very good correlation (above 80%), and capture the drift completely. This shows again that CTAs are simply following a long-term trending signal, and there is little added value in their idiosyncrasies.

However, this also shows us that a CTA does not provide the same hedge a single-asset trend provides: some of the convexity is lost because of diversification. We however have found that CTAs do offer an interesting hedge to Risk-Parity products, which we approximated with a very good precision by long positions on the main asset classes. This property is quite new, and we feel it makes the trend a valid addition in the book of any manager holding Risk Parity products (or simply a diversified long position in both equities and bonds).

We then turned our attention to the link between trend and volatility. We found that a simple trending toy-model shares an exposure to the long-term variance with a naked straddle. The difference is the fact that the entry price for the straddle is fixed by the at-the-money volatility, while the trend pays the realized short-term variance. We then propose a very clean way to get exposure to this short term variance by using the trending toy-model as a hedging strategy for a portfolio of strangles. This is a simple, model-free portfolio that offers the same pay-off than traditional variance swaps.

All in all, these results prove that a trending system does offer protection to long-term large moves of the market. This, coupled with the high significance of this market anomaly, really sets it apart in the world of investment strategies. A more pressing issue could be the capacity of this strategy, but recent performance seems to be quite in line with long-term returns, so there is little evidence of over-crowding.

The future of CTA industry is still unclear, whether large established funds will continue to charge their usual fees, or if cheaper investment vehicles will attract interest. What we have seen in this paper is that there is little value in the "superior skill" of a single CTA, unless this CTA has correlation with the SG CTA Index that is significantly below 80% (i.e. it does something different from what we have proposed). For a simple trend exposure (which is desirable in a diversified portfolio), a cheap alternative may well be enough.
6 Appendix

6.1 Proof of trend formula

6.1.1 Theorem

Let $X_t$ be a discrete process and $F_\alpha$ be the following linear filter:

$$F_\alpha[X_t] = \sum_{i \geq 0} \alpha^i X_{t-i} \quad (\alpha > 0) \quad (10)$$

we have the following relation:

$$F_\alpha[F_\alpha[Y_t] + X_t F_\beta[Y_t]] = F_\alpha[X_t] F_\beta[Y_t] + F_\alpha F_\beta[X_t Y_t] \quad (11)$$

**Proof**

Using the definition of the filter, we decompose the first term of the left hand-side:

$$F_\alpha[X_t] F_\beta[Y_t] = \sum_{i \geq 0} \alpha^i \beta^j X_{t-i} Y_{t-j} + \sum_{i \geq 0} \sum_{j > i} \alpha^i \beta^j X_{t-i} Y_{t-j} + \sum_{i \geq 0} \sum_{j > i} \alpha^i \beta^j X_{t-i} Y_{t-j} \quad (12)$$

The diagonal term:

$$\sum_{i \geq 0} \sum_{j > i} \alpha^i \beta^j X_{t-i} Y_{t-j} = \sum_{i \geq 0} \sum_{j > i} \alpha^i \beta^j X_{t-i} \beta^{j-i} Y_{t-i-(j-i)} = \sum_{i \geq 0} \sum_{j > i} \alpha^i \beta^j X_{t-i} \beta^{j-i} Y_{t-i-k} \quad \text{by taking } k = j - i$$

$$= \sum_{i \geq 0} \sum_{j > i} \alpha^i \beta^j (X_{t-i} F_\beta[Y_{t-i}] - Y_{t-i})$$

$$= \sum_{i \geq 0} \sum_{j > i} \alpha^i \beta^j (X_{t-i} F_\beta[Y_{t-i}] - \sum_{i \geq 0} \alpha^i \beta^j X_{t-i} Y_{t-i})$$

$$= \sum_{i \geq 0} \sum_{j > i} \alpha^i \beta^j U_{t-i} - \sum_{i \geq 0} \sum_{j > i} \alpha^i \beta^j X_{t-i} Y_{t-i} \quad \text{by taking } U_t = X_t F_\beta[Y_t]$$

$$= F_{\alpha \beta}[U_t] - F_{\alpha \beta}[X_t Y_t] = F_{\alpha \beta}[X_t Y_t] - F_{\alpha \beta}[X_t Y_t] \quad (13)$$

Apply the same trick to the other cross term, we have:

$$\sum_{i \geq 0} \sum_{j > i} \alpha^i \beta^j X_{t-i} Y_{t-j} = F_{\alpha \beta}[Y_t F_\alpha[X_t]] - F_{\alpha \beta}[X_t Y_t] \quad (14)$$

The diagonal term:

$$\sum_{i \geq 0} \alpha^i \beta^j X_{t-i} Y_{t-i} = F_{\alpha \beta}[X_t Y_t] \quad (15)$$

Using (12), (13), (14) and (15), we have:

$$F_\alpha[X_t] F_\beta[Y_t] = F_{\alpha \beta}[X_t Y_t] + \left(F_{\alpha \beta}[X_t F_\beta[Y_t]] - F_{\alpha \beta}[X_t Y_t]\right) + \left(F_{\alpha \beta}[Y_t F_\alpha[X_t]] - F_{\alpha \beta}[X_t Y_t]\right)$$

which gives:

$$F_{\alpha \beta}[Y_t F_\alpha[X_t] + X_t F_\beta[Y_t]] = F_{\alpha \beta}[X_t] F_\beta[Y_t] + F_{\alpha \beta}[X_t Y_t] \quad (16)$$

We remark that $F_\alpha[X_t] = X_t + \alpha F_\alpha[X_{t-1}]$, we can rewrite the above formula as:

$$F_{\alpha \beta}[\alpha Y_t F_\alpha[X_{t-1}] + \beta X_t F_\beta[Y_{t-1}]] = F_{\alpha \beta}[X_t] F_\beta[Y_t] - F_{\alpha \beta}[X_t Y_t] \quad (17)$$
6.1.2 Discrete equation for EMA filter

We now employ the above result to derive the main result given in Section 2 for an exponential moving average filter. Taking \( \alpha = 1 - 2/(\tau + 1) \), and renormalize the filter \( \mathcal{F}_\alpha \) in a way that the sum of weights is equal to 1. Let \( \mathcal{L}_r \) be the renormalized filter:

\[
\mathcal{L}_r \equiv (1 - \alpha)\mathcal{F}_\alpha
\]

The timescales related to the parameter \( \alpha \) and \( \alpha^2 \) are given respectively by:

\[
\tau = \frac{1 + \alpha}{1 - \alpha}, \quad \tau' = \frac{1 + \alpha^2}{1 - \alpha^2} = \frac{\tau}{2} + \frac{1}{2\tau}
\]

We now rewrite the above result with the new filter \( \mathcal{L}_r \) and the parameter \( \tau \) instead of \( \mathcal{F}_\alpha \) and parameter \( \alpha \). \( \mathcal{F}_\alpha = (\tau + 1)\mathcal{L}_r/2 \), \( \mathcal{F}_\alpha^2 = (\tau + 1)^2\mathcal{L}_r'/4\tau \). Insert these expressions into the result of the above theorem, we obtain the equation for discrete EMA filter:

\[
(1 - \frac{1}{\tau})\mathcal{L}_r' \left[ X_t\mathcal{L}_r[X_{t-1}] \right] = \left( \mathcal{L}_r[X_t] \right)^2 - \frac{1}{\tau} \mathcal{L}_r'[X_t^2]
\]

6.2 Generalized trend formula for trend-following

6.2.1 Trend formula in continuous-time framework

In order to derive the generalized formula, we employ the continuous-time approach framework. Within this framework, we consider the stochastic process of the asset price \( S_t \). In order to derive the generalized form, we employ the continuous-time approach framework.

Writing the above equation in form of a non-linear function \( P_t \)

\[
\phi \equiv \frac{d}{dt} F(P_t)
\]

In this form, the above filter can be also interpreted as the exponential moving average filter \( \mathcal{L}_r \) in the discrete-time framework. Hence, we employ the notation \( \mathcal{L}_r \) to rewrite the above filter equation \( P_t = \mathcal{L}_r[S_t] \). Building the trend-following strategy using this trend estimation deformed by a non-linear function \( \phi(x) \), we obtain the of the change of profit and loss (\( G_t \)):

\[
dG_t = \phi(P_t) \times dS_t.
\]

Here \( \phi(x) \) can be a function like \( \phi(x) = \text{sign}(x) \) or \( \phi(x) = \text{Cap}_{\lambda_1, \lambda_2}(x) \) in order to limit the extreme exposure to the asset.

Using the Kalman filter (Eq. [20]) to eliminate the dependence of the change of profit and loss (Eq. [22]) on the asset price \( S_t \), we obtain:

\[
dG_t = \phi(P_t)P_t dt + \frac{\tau}{2} \phi(P_t) dP_t
\]

Let \( F(x) \) be the primitive of \( \phi(x) \), then using Itô’s lemma we have:

\[
dF(P_t) = \phi(P_t) dP_t + \frac{2\phi'(P_t)}{\tau} dS_t^2
\]

Insert this expression in the equation of \( PnL_t \), we obtain:

\[
dG_t = \frac{\tau}{2} dF(P_t) + \left[ \phi(P_t) P_t dt - \frac{\phi'(P_t)}{\tau} dS_t^2 \right]
\]

We rearrange different terms of the \( PnL \) equation and introduce new timescale \( T \):

\[
d\left( \frac{T}{\tau} F(P_t) \right) = \frac{\tau}{T} \left( \frac{T}{\tau} F(P_t) \right) dt + \frac{2}{T} \left[ dG_t + \frac{\phi'(P_t)}{\tau} dS_t^2 - \phi(P_t) P_t dt + \frac{T}{\tau} F(P_t) dt \right]
\]

Writing the above equation in form of \( \mathcal{L}_\tau \) filter, we obtain:

\[
\frac{T}{\tau} F(P_t) = \mathcal{L}_\tau \left[ dG_t + \frac{\phi'(P_t)}{\tau} dS_t^2 - \phi(P_t) P_t dt + \frac{T}{\tau} F(P_t) dT \right]
\]

Finally, we derive the generalized equation for trend \( PnL \):

\[
\mathcal{L}_\tau[dG_t] = \frac{T}{\tau} F(P_t) - \mathcal{L}_\tau \left[ \frac{\phi'(P_t)}{\tau} dS_t^2 \right] + \mathcal{L}_\tau \left[ \phi(P_t) P_t - \frac{T}{\tau} F(P_t) \right] dt
\]
6.2.2 Linear trend estimation

In the case where φ(x) = x, we have F(x) = x^2/2, then we find the result showed for discrete approach:

\[ \mathcal{L}_r[dG] = \frac{\tau}{2T}P_t^2 - \frac{1}{\tau} \mathcal{L}_r[dS_t^2] + \left(1 - \frac{\tau}{2T}\right) \mathcal{L}_r[P_t^2]dt \]  \quad (24)

With the choice of timescale \( T = \tau/2 \) we eliminate the last term (correction term) then obtain:

\[ \mathcal{L}_r[dG] = P_t^2 - \frac{1}{2T} \mathcal{L}_r[dS_t^2] \]  \quad (25)

6.2.3 Non-linear trend estimation

Let us consider now the case \( \phi(x) = \text{sign}(x) \) hence its primitive is \( F(x) = |x| \) and its derivative is \( \phi'(x) = 2\delta(x) \). We obtain:

\[ \mathcal{L}_r[dG] = \frac{\tau}{T} |P_t| - \frac{2}{\tau} \mathcal{L}_r \left[ \delta(P_t) dS_t^2 \right] + \left(1 - \frac{\tau}{T}\right) \mathcal{L}_r[|P_t|]dt \]

With the choice of timescale \( T = \tau \), we eliminate the last term (correction term) then obtain the following equation:

\[ \mathcal{L}_r[dG] = |P_t| - \frac{2}{\tau} \mathcal{L}_r \left[ \delta(\bar{\mu}_t) dS_t^2 \right] \]

For return \( dS_t \) is risk managed at stable volatility \( \sigma \) and follows Gaussian process \( dS_t \sim \mathcal{N}(\bar{\mu}, \sigma) \), we have the following approximation:

\[ \langle \mathcal{L}_r \left[ \delta(P_t) dS_t^2 \right] \rangle_P \approx \sigma^2 \langle \delta(P) \rangle \]

As the trend estimate \( P_t \) follows the following distribution \( \mathcal{N}(\bar{\mu}, \sigma/\sqrt{T}) \), we have:

\[ \langle \delta(P) \rangle = \int_{-\infty}^{\infty} \delta(x) \frac{1}{\sqrt{2\pi}\sigma_x} e^{-x^2/(2\sigma_x^2)} dx = \sqrt{\frac{\tau}{2\pi}} \frac{1}{\sigma} \]

Insert this approximation in the PnL equation, we obtain the result for the case of sign of trend:

\[ \langle \mathcal{L}_r[dG] \rangle_P = |P_t| - \sqrt{\frac{2}{\pi\sigma}} \]  \quad (26)

Renormalize the predictor as \( P_t \rightarrow P_t/\sqrt{T}/\sigma \), we obtain the formula:

\[ \langle \mathcal{M}_r[dG] \rangle_P = \sqrt{T} \left( |P_t| - \sqrt{\frac{2}{\pi}} \right) \]  \quad (27)

6.3 Effect of diversification

We provide here the proofs of the convexity of the trend with respect to the index and the risk-parity portfolio for the multi-asset case.

6.3.1 Convexity versus index

Let us consider a multi-asset trend following portfolio with weight distribution \( \{w_k\} \) for \( k = 1 \ldots N \) over \( N \) assets with return \( \Delta_k \). We derive now convexity with respect to a given asset (S&P index for example) denoted by index \( s \). Let us define some notation:

\[ \bar{P}_{k,t} = P_{k,t} - \alpha_k \quad \text{with} \quad \alpha_k = \mathbb{E}[P_{k,t}] \quad \text{and} \quad \nu_k^2 = \mathbb{E}[\bar{P}_{k,t}^2] \]

Therefore \( P_{s,t} \) is the trend estimated on the return of the asset \( s \). We project now the individual trend \( P_{k,t} \) on this direction and obtain:

\[ P_{k,t} = \beta_k \bar{P}_{s,t} + \alpha_{k,t} \quad \text{with} \quad \beta_k = \frac{\mathbb{E}[\bar{P}_{k,t} \bar{P}_{s,t}]}{\nu_s^2} \]

Insert this decomposition in the sum of individual trend formula and perform some algebraic calculations, we obtain finally:

\[ \langle \mathcal{M}_r[G_t] \mid P_{s,t} = p \rangle \rightarrow \frac{1}{\beta_s^2} \left( p^2 - \nu_s^2 + \alpha_s^2 \right) \]
in which the different average quantities $\beta_2$, $\alpha_2$, $\bar{\alpha}$, $\hat{\nu}_2$ are defined as below:

$$
\beta_2 = \frac{1}{N} \sum_{k=1}^{N} \beta_k^2, \quad \alpha_2 = \frac{\sum_{k=1}^{N} \alpha_k^2}{\sum_{k=1}^{N} \beta_k^2}, \quad \bar{\alpha} = \frac{\sum_{k=1}^{N} \beta_k \alpha_k}{\sum_{k=1}^{N} \beta_k^2}, \quad \hat{\nu}_2 = \mathbb{E}[\hat{P}_{s,t}^2] = \nu_2^2 + \alpha^2
$$

Note that if all drifts are zero in average and if the time series are properly normalized the above equation simply writes:

$$
\mathcal{M}_t \left[ \tilde{G}_t \right] \approx \beta_2 \left( \hat{P}_{s,t}^2 - 1 \right)
$$

### 6.3.2 Convexity versus risk parity

We can define the total trend indicator:

$$
\tilde{P}_t = \sum_{k=1}^{N} w_k P_{k,t}
$$

Applying the trend formula for individual products then performing the weighted sum using $\{w_k\}$, we obtain:

$$
\mathcal{M}_t \left[ \tilde{G}_t \right] \geq l \left( \tilde{P}_t^2 - \sum_{k=1}^{N} w_k \mathcal{L}_t \left[ \tilde{\Delta}_{k,t}^2 \right] \right)
$$

Taking conditional expectation on the total trend, we obtain similar result for the single asset case:

$$
\left( \mathcal{M}_t \left[ \tilde{G}_t \right] | \tilde{P}_t = p \right) \geq l \left( \tilde{P}_t^2 - \sum_{k=1}^{N} w_k \mathcal{L}_t \left[ \tilde{\Delta}_{k,t}^2 \right] \right) | \tilde{P}_t = p \rightarrow l (p^2 - 1)
$$

This result provides a lower bound of the convexity with respect the weighted average trend.

We now rewrite the above inequality in form of the filter $\mathcal{M}_t$ instead of the EMA-filter $\mathcal{L}_t$ for both sides then obtain:

$$
\mathcal{M}_t \left[ \tilde{G}_t \right] \geq l \left( \frac{\mathcal{M}_t \left[ R_t \right]^2}{\tau} - 1 \right)
$$

### 6.4 Strangle portfolio: Mark-to-the-market and Greeks

In this appendix, we compute the mark-to-the-market PnL of the above payoff:

$$
PnL_{[0,T]} = R_{[0,T]}^2 - T \hat{\sigma}_T^2
$$

Here, $\hat{\sigma}_T^2$ is the implied variance of daily return.

$$
\hat{\sigma}_T^2 = \frac{2}{T} \left( \int_{S_0}^{S_T} P_{K,T} + \int_{S_0}^{S_T} C_{K,T} \right)
$$

Let us employ the following decomposition at time $t \in [0,T]$:

$$
R_{[0,T]}^2 = \left( R_{[0,t-1]} + R_{[t,T]} \right)^2
= R_{[0,t-1]}^2 + 2R_{[0,t-1]} R_{[t,T]} + R_{[t,T]}^2
$$

Pricing this quantity at time $t$ by the market-measure $\mathbb{E}$ gives us the mark-to-the-market value of this payoff:

$$
PnL_{[0,t]} = \mathbb{E}_t \left( R_{[0,t]}^2 \right) - T \hat{\sigma}_T^2
= R_{[0,t-1]}^2 + R_{[0,t-1]} \mathbb{E}_t \left( R_{[t,T]} \right) + \mathbb{E}_t \left( R_{[0,t]}^2 \right) - T \hat{\sigma}_T^2
= R_{[0,t-1]}^2 + R_{[0,t-1]} \mathbb{E}_t \left( S_T - S_t \right) + \mathbb{E}_t \left( R_{[t,T]}^2 \right) - T \hat{\sigma}_T^2
$$

19
Here, we have to price a new portfolio of strangles \( E_t \left( R_{[t,T]}^2 \right) \) centered at \( S_t \) and a future contract \( E_t \left( S_T - S_t \right) \) with fair price zero. The new strangles portfolio gives us simply \( \tau \bar{\sigma}^2 \) while the fair price of future contract is zero. Hence, we obtain finally the market price of our original portfolio:

\[
PnL_{[0,t]} = R_{[0,t-1]}^2 + \tau \bar{\sigma}^2 - T \bar{\sigma}^2
\]

Here \( \tau = T - t \) is the time-to-maturity. Therefore, the daily mark-to-the-market PnL is given by:

\[
PnL_t = PnL_{[0,t+1]} - PnL_{[0,t]} = R_t^2 + \left( \left( (\tau - 1) \bar{\sigma}_{t-1}^2 - \tau \bar{\sigma}_t^2 \right) + 2 R_{[0,t-1]} R_t \right)\]

\[
= \left( R_t^2 - \bar{\sigma}_{t-1}^2 \right) + \tau \left( \bar{\sigma}_{t-1}^2 - \bar{\sigma}_t^2 \right) + 2 (S_t - S_0) R_t
\]

From the above equation, we deduce simply the global Greeks of the strangle portfolio:

\[
\Gamma = 1, \quad \mathcal{V} = 2 \tau \bar{\sigma}, \quad \Delta_t = 2 (S_t - S_0)
\]
References